

# Testes não paramétricos duas amostras independentes

Professor: Pedro M. Almeida-Junior

---

29 de julho de 2021

Departamento de Estatística (UEPB)

# O Teste Exato de Fisher

- O teste exato de Fisher é uma técnica extremamente útil para analisar dados discretos (nominal ou ordinal) quando as duas amostras independentes são pequenas.
- Ele é usado quando todas as observações de duas amostras aleatórias independentes caem em uma ou em outra classe entre duas classes mutuamente exclusivas.
- Os escores são representados por frequências em uma tabela de contingência  $2 \times 2$ .

**Tabela 1:** Tabela de contingência  $2 \times 2$

Variável	Grupo		Total linha
	I	II	
+	$A$	$B$	$A + B$
−	$C$	$D$	$C + D$
Total coluna	$A + C$	$B + D$	$N$

- Os Grupos I e II podem ser dois grupos independentes quaisquer, tais como experimentais e controles, homens e mulheres, etc ...
- Os nomes das linhas, aqui arbitrariamente indicados como mais (+) e menos (−) podem ser duas classificações quaisquer: aprovado e reprovado, concorda e discorda, etc.
- O teste determina se os dois grupos diferem na proporção com a qual eles caem nas duas classificações.

# Exigências do Teste

- ✓ Amostras aleatórias e independentes;
- ✓ Duas classes mutuamente exclusivas;
- ✓ Nível de Mensuração em escala nominal ao menos.
- ✓ Totais das linhas e colunas devem ser fixados antes do experimento.

## O método

A probabilidade exata de observar um conjunto particular de frequências em uma tabela  $2 \times 2$ , quando os totais marginais são considerados como fixos, é dada pela distribuição hipergeométrica:

$$\begin{aligned} p &= \frac{\binom{A+C}{A} \binom{B+D}{B}}{\binom{N}{A+B}} \\ &= \frac{[(A+C)!/A!C!][B+D)!/B!D!]}{N!/[(A+B)!(C+D)!]} \end{aligned}$$

e então,

$$p = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!} \quad (1)$$

## Exemplo

**Tabela 2:** Tabela de contingência  $2 \times 2$

Variável	Grupo		Total
	I	II	
+	5	4	9
−	0	10	10
Total	5	14	19

Então, pela equação 1 o valor de  $p$  é dado por

$$\begin{aligned} p &= \frac{9!10!5!14!}{19!5!4!0!10!} \\ &= 0,0108 \end{aligned} \tag{2}$$

A seguir estão as hipóteses nulas que podem ser testadas e suas alternativas.

### 1. (Bilateral)

- $H_0$  : A proporção com a característica de interesse é a mesma em ambas as populações; ou seja,  $p_1 = p_2$ .
- $H_1$  : A proporção com a característica de interesse não é a mesma nas duas populações;  $p_1 \neq p_2$ .

### 2. (Unilateral)

- $H_0$  : A proporção com a característica de interesse na população 1 é menor ou igual à proporção na população 2;  $p_1 \leq p_2$ .
- $H_1$  : A proporção com a característica de interesse é maior na população 1 do que na população 2;  $p_1 > p_2$ .

## Cálculo do p-valor

- O teste exato de Fisher consiste no cálculo da probabilidade  $p$  e das configurações mais extrema, a que sugere a rejeição de  $H_0$ ;

$$\text{valor-}p = p + p_k + \cdots + p_0$$

seja  $k$  a quantidade de tabelas possíveis, temos que  $p_k + \cdots + p_0$  são as probabilidades das configurações mais extremas à rejeição de  $H_0$ .

- As possíveis tabelas que, com os mesmos totais marginais fixos, indicam ainda mais desvio em relação à hipótese nula. A seguir apresentamos uma ilustração para obtenção das tabelas mais extremas a hipótese nula.



# Tabelas com afastamento da Hipótese nula

$a - 1$	$b + 1$	$a + b$
$c + 1$	$d - 1$	$c + d$
$a + c$	$b + d$	$n$

$a - 2$	$b + 2$	$a + b$
$c + 2$	$d - 2$	$c + d$
$a + c$	$b + d$	$n$

0	$a + b$	$a + b$
$c + a$	$d - a$	$c + d$
$a + c$	$b + d$	$n$

## Exemplo

**Exemplo:** O objetivo do estudo foi avaliar a eficácia a longo prazo de tomar *indinavir/ritonavir* duas vezes ao dia em combinação com dois inibidores da transcriptase reversa de nucleosídeos entre indivíduos HIV-positivos que foram divididos em dois grupos. O Grupo 1 consistiu em pacientes que não tinham histórico de uso de inibidores da protease (PI Naive). O Grupo 2 consistia em pacientes com história prévia de uso de inibidor de protease (PI Experiente). A Tabela seguinte apresentamos os indivíduos que permaneceram no regime por 120 semanas de acompanhamento. **Queremos saber se podemos concluir que os pacientes classificados como grupo 1 têm uma probabilidade menor do que os indivíduos do grupo 2 de permanecer no regime por 120 semanas.**

## Tabela de contingência $2 \times 2$

Grupos	Permanecer no regime por mais de 120 semanas		Total
	Sim	Não	
Grupo 1: Naive	2	7	9
Grupo 2: Experiente	8	4	12
Total	10	11	21

- **Hipóteses:**

- $H_0$ : a proporção de indivíduos que permaneceram 120 semanas no regime classificados como grupo 2 é igual ou menor que a proporção de indivíduos classificados como grupo 1. ( $p_2 \leq p_1$ )
- $H_1$  : Os pacientes do Grupo 2 têm uma taxa maior do que os pacientes do Grupo 1 de permanecer no regime por 120 semanas. ( $p_2 > p_1$ )

Estes são os passos no uso do teste exato de Fisher:

1. Coloque as frequências observadas em uma tabela  $2 \times 2$ .
2. Determine os totais marginais (Totais de linhas).
3. Determine o valor-p (unilateral ou bilateral)
4. Tome uma decisão baseada no valor-p e interprete os resultados.

**Exemplo:** Um pesquisador realizou um ensaio clínico para investigar a relação entre o tempo de sono profundo e dois grupos de idades. Com base na Tabela abaixo, podemos afirmar se as pessoas que tem idade entre 30 e 45 anos possuem mais tempo de sono profundo que as pessoas entre 65 e 80 anos ?

Idade	Tempo de sono profundo		Total
	Mais de 3 horas	Menos de 3 horas	
30 a 45 anos	7	4	11
65 a 80 anos	3	8	11
Total	10	12	22

# Teste Exato de Fisher

- O Teste Exato de Fisher é a alternativa ao caso de duas amostras independentes, quando o tamanho da amostra é pequeno, pois nesse caso o teste  $\chi^2$  não se aplica.
- Apropriado para variáveis dicotômicas e em escala nominal.